AMENDMENTS TO THE CLAIMS

1. (Previously Presented) A method for classifying an audio signal containing speech information, the method comprising:

receiving the audio signal;

classifying a sound in the audio signal as a vowel class when a first phoneme-based model determines that the sound corresponds to a sound represented by a set of phonemes that define vowels;

classifying the sound in the audio signal as a fricative class when a second phoneme-based model determines that the sound corresponds to a sound represented by a set of phonemes that define consonants; and

classifying the sound in the audio signal based on at least one non-phoneme based model, the at least one non-phoneme based model including at least one model for classifying the sound in the audio signal based on bandwidth.

2. (Previously Presented) The method of claim 1, wherein the at least one non-phoneme based model includes models for classifying the sound in the audio signal based on speaker gender.

3. (Original) The method of claim 1, wherein the at least one non-phoneme based model includes a model for classifying the sound in the audio signal as silence.

4. (Original) The method of claim 1, further comprising:

initially converting the audio signal into a frequency domain signal.

5. (Original) The method of claim 1, further comprising:

generating cepstral features for the audio signal.

6. (Original) The method of claim 1, wherein the fricative class includes phonemes that relate to fricatives and obstruents.

7. (Original) The method of claim 1, wherein the first and second phoneme-based models are Hidden Markov Models.

8. (Original) The method of claim 1, further comprising:

classifying the sound in the audio signal as a coughing class when the sound corresponds to a non-speech sound.

9. (Original) The method of claim 8, wherein the non-speech sound includes at least one of coughing, laughter, breath, and lip-smack.

10. (Previously Presented) A method of training audio classification models, the method comprising:

receiving a training audio signal;

receiving phoneme classes corresponding to the training audio signal;

training a first Hidden Markov Model (HMM), based on the training audio signal and the phoneme classes, to classify speech as belonging to a vowel class when the first HMM determines that the speech corresponds to a sound represented by a set of phonemes that define vowels;

training a second HMM, based on the training audio signal and the phoneme classes, to classify speech as belonging to a fricative class when the second HMM determines that the speech corresponds to a sound represented by a set of phonemes that define consonants; and

training at least one model to classify the sound based on a bandwidth of the sound.

11. (Original) The method of claim 10, wherein the phoneme classes include information that defines word boundaries.

12. (Original) The method of claim 11, wherein the method further comprises:

receiving a sequence of transcribed words corresponding to the audio signal; and

generating the information that defines the word boundaries based on the transcribed words.

13. (Canceled)

14. (Original) The method of claim 10, further comprising:

training at least one model to classify the sound based on gender of a speaker of the sound.

15. (Original) The method of claim 10, wherein the fricative class includes phonemes that relate to fricatives and obstruents.

16. (Previously Presented) An audio classification device comprising:

a signal analysis component configured to receive an audio signal and process the audio signal by at least one of the converting the audio signal to the frequency domain and generating cepstral features for the audio signal; and

a decoder configured to classify portions of the audio signal as belonging to at least one of the plurality of classes, the classes including

a first phoneme-based class that applies to the audio signal when a portion of the audio signal corresponds to a sound represented by a set of phonemes that define vowels,

a second phoneme-based class that applies to the audio signal when a portion of the audio signal corresponds to a sound represented by a set of phonemes that define consonants, and

at least one non-phoneme class;

wherein the decoder determines the at least one non-phoneme class using models that classify the portions of the audio signal based on bandwidth.

17. (Original) The audio classification device of claim 16, wherein the second phoneme-based class includes fricative phonemes and obstruent phonemes.

18. (Original) The audio classification device of claim 16, wherein the first and second phoneme-based classes are determined based on hidden Markov Models.

19. (Previously Presented) The audio classification device of claim 16, wherein the decoder determines the at least one non-phoneme class using models that classify the portions of the audio signal based on speaker gender.

20. (Original) The audio classification device of claim 16, wherein the decoder determines the at least one non-phoneme class using a model that classifies the portions of the audio signal as silence.

21. (Original) The audio classification device of claim 16, wherein the plurality of classes additionally include:

a third phoneme-based class that applies to the audio signal when a portion of the audio signal corresponds to a non-speech sound.

22. (Original) The audio classification device of claim 21, wherein the non-speech sound includes at least one of the coughing, laughter, breath, and lip-smack.

23. (Original) A system comprising:

an indexer configured to receive input audio data and generate a rich transcription from the audio data, the indexer including:

audio classification logic configured to classify the input audio data into at least one of a plurality of broad audio classes, the broad audio classes including a phoneme-based

vowel class, a phoneme-based fricative class, a non-phoneme based bandwidth class, and a non-phoneme based gender class,

a speech recognition component configured to generate the rich transcription based on the broad audio classes determined by the audio classification logic;

a memory system for storing the rich transcription; and

a server configured to receive requests for documents and respond to the requests by transmitting one or more of the rich transcriptions that match the requests.

24. (Original) The system of claim 23, wherein the broad audio classes further include a phoneme-based coughing class.

25. (Original) The system of claim 24, wherein the coughing class includes sounds relating to coughing, laughter, breath, and lip-smack.

26. (Original) The system of claim 23, wherein the phoneme-based fricative class includes phonemes that define fricative or obstruent sounds.

27. (Original) The system of claim 23, wherein the indexer further includes at least one of: a speaker clustering component, a speaker identification component, a name spotting component, and a topic classification component.

28. (Previously Presented) A device comprising:

means for classifying a sound in an audio signal as a vowel class when a first phoneme-based model determines that the sound corresponds to a sound represented by a set of phonemes that define vowels;

means for classifying the sound in the audio signal as a fricative class when a second phoneme-based model determines that the sound corresponds to a sound represented by a set of phonemes that define consonants; and

means for classifying the sound in the audio signal based on at least one non-phoneme based model, the at least one non-phoneme based model including at least one model for classifying the sound in the audio signal based on bandwidth.

29. (Original) The device of claim 28, further comprising:
means for converting the audio signal into a frequency domain signal.

30. (Original) The device of claim 28, further comprising:
means for generating cepstral features for the audio signal.

31. (Original) The device of claim 28, further comprising:
means for classifying the sound in the audio signal as a coughing class when the sound corresponds to a non-speech sound.